

PROJECT TITLE	Hanen Learning Language and Loving It (Hanen LLLI)
DEVELOPER (INSTITUTION)	The Hanen Centre
EVALUATOR (INSTITUTION)	National Centre for Social Research
PRINCIPAL INVESTIGATOR(S)	Jonah Bury
PROTOCOL AUTHOR(S)	Sashka Dimova, Molly Scott, Berenice Scandone, Alessandra Sciarra, Mansor Rezaian, Jonah Bury
TRIAL DESIGN	Two-arm cluster randomised controlled trial with random allocation at Early Years setting level
TRIAL TYPE	Efficacy
PUPIL AGE RANGE AND KEY STAGE	3 to 4 years old, Early years
NUMBER OF NURSERIES	165
NUMBER OF PUPILS	2,805
PRIMARY OUTCOME MEASURE AND SOURCE	Receptive language measured with the British Picture Vocabulary Scale (BPVS), Third Edition, GL Assessment.
SECONDARY OUTCOME MEASURES AND SOURCE	<ol style="list-style-type: none"> Expressive and receptive language measured with the Renfrew Action Picture Test (RAPT), 5th Edition, Routledge Socio-emotional development measured with the Strength and Difficulties Questionnaire - Teacher version (SDQ-T), Robert N. Goodman

Protocol version history

VERSION	DATE	REASON FOR REVISION
1.0 [<i>original</i>]	14 July, 2022	N/A

Table of Contents

Protocol version history	1
Study rationale and background	2
Intervention.....	4
Impact evaluation	7
<i>Research questions</i>	7
<i>Design</i>	8
<i>Randomisation</i>	9
<i>Participants</i>	10
<i>Sample size calculations</i>	11
<i>Outcome measures</i>	13
<i>Analysis</i>	15
<i>Compliance</i>	16
<i>Longitudinal follow-ups</i>	17
Implementation and process evaluation	17
<i>Research questions</i>	18
<i>Research methods</i>	18
<i>Analysis</i>	22
Cost evaluation	23
Ethics and registration.....	23
Data protection	23
Personnel	24
Risks.....	25
Timeline.....	26
References	28
Appendix 1: Intervention logic model	31

Study rationale and background

Policy background

Early language skills are a crucial building block for children’s development. Children naturally develop language skills at different rates, but some children fall behind at an early age. Depending on the measures used, only 73% of children were found to have reached the expected level across the Communication and Language and Literacy domains of the Early Years Foundation Stage (EYFS) Profile in 2019 (DfE, 2019a). Children from socially disadvantaged backgrounds (indicated by free school meals eligibility) are, on average, much more likely to experience delays and difficulties in their language development.

The importance of language skills is reflected in the revised [EYFS Framework](#), which became statutory at the beginning of the academic year 2021/22 (DfE, 2021b). Reforms to the [EYFS](#)

[Early Learning Goals \(ELGs\)](#)¹ (DfE, 2021c) were implemented in response to a review of Reception year in 2017 which recommended greater focus on spoken language and vocabulary development (Ofsted, 2017).

A key determinant of language development is the amount and quality of language to which a young child is exposed. Given that over 95% of 3-4-year-olds participate in formal early education in England (DfE, 2018), Early Years (EY) nursery staff play an important role. Although Ofsted assessed the majority (96%) of EY settings as good or outstanding (Ofsted, 2020), the EY workforce is comprised predominantly of Level 3 qualified staff (below degree level) (DfE, 2019b) and recent research points to a downward trend in qualifications (EPI, 2018, 2020). It is likely therefore, that the workforce would benefit from continuing professional development (CPD) targeted at language.

Existing evidence

A review of language interventions by Law *et al.* (2017) recommended further research on the effectiveness of training for nursery staff to deliver programmes within EY settings. The review reported on three evaluations of Hanen LLLI. An exploratory study by Girolametto *et al.* (2003) in which 16 US teachers of 3 to 5-year-olds were randomly assigned to receive training in Hanen LLLI found that pupils who engaged in shared reading and playdough activities used a greater number of utterances, multiword combinations, and peer directed utterances. However, the number of different words did not differ by group. Cabell *et al.* (2011) reported on a Randomised Controlled Trial (RCT) of a programme that trained 49 teachers of 3 to 5-year-old kindergarteners in the US in Hanen LLLI. The evaluation found that children assigned to the treatment group demonstrated improvement in expressive vocabulary, but there was no difference between the treatment and control groups on spoken language outcomes. Piasta *et al.* (2012) reported on an RCT of a US programme based on Hanen LLLI in which 49 preschool EY practitioners received training. The authors found a positive difference in the total utterances, number of different words used, and mean length of utterances. Overall, the review found that Hanen LLLI was particularly promising with high effect sizes (albeit based on low security of findings) for the impact of professional development on nursery staff's conversational responsivity and children's linguistic productivity and complexity (Law *et al.*, 2017). The authors concluded that most language interventions focus on improving vocabulary, whereas Hanen LLLI recognised the importance of conversation and oral narrative.

A handful of Hanen LLLI training programs have been run in the UK in the last few years, focusing mainly on a shortened version of Hanen LLLI, called Teacher Talk. However, no impact evaluations of Hanen LLLI have taken place in the UK to date. A small-scale [pilot evaluation of Hanen LLLI](#) conducted by NatCen Social Research in 2019 explored evidence of promise (e.g. do providers perceive the programme to have impact on children's language?), feasibility (e.g. was the programme delivered as intended?) and readiness for trial (e.g. what changes, if any, are needed to the intervention theory?). The pilot found the intervention to be attractive to EY settings and showed evidence of promise regarding changes to nursery staff's interactions with children. Based on observations of workshop, video feedback sessions and interviews with nursery managers, nursery staff, and speech and language therapists, the pilot found that delivery was largely as intended. However, there were some deviations, for instance in the duration of video feedback sessions (15 minutes instead of 30 minutes). Nursery staff were engaged, and attendance of the workshops and video

¹ ELGs summarise the knowledge, skills and understanding that all young children should have gained by the end of the academic year in which they turn 5, the reception year.

feedback sessions was high. Program Leaders and nursery staff reported that the video feedback sessions were key in boosting engagement as nursery staff could see how they were progressing and where they could still improve. Interviews with nursery staff in settings² indicate that practitioners who were already using strategies in line with Hanen LLLI did not perceive impacts to their daily practice. However, these practitioners did feel that Hanen LLLI made them more aware of how they were interacting with children and therefore more focused in their interactions. In contrast, those who were not already using Hanen LLLI strategies perceived a more significant impact on their day-to-day practice, such as providing more opportunities for children to initiate conversations. Nursery staff also reported improvements in children's language and communication development, such as an increase in expressive vocabulary and listening skills. Improvements were perceived across all pupil groups, including pupils who speak English as an additional language (EAL) and those with special educational needs (SEN).

Following the pilot, which demonstrated that the programme is feasible in the English context, Hanen LLLI was recommended for further evaluation. NatCen was commissioned to carry out the efficacy trial of Hanen LLLI following the pilot.³ The initial efficacy trial was cancelled in March 2021 due to the perceived risks and ethical challenges of continuing with outcome testing during the COVID-19 pandemic. In the remainder of this protocol, we refer to the cancelled trial of Hanen LLLI as Hanen LLLI 1. A standalone IPE was carried out for Hanen LLLI 1, which will be published in Summer 2022 and confirms the attractiveness of the intervention for EY settings and perceived benefits for children and staff.

This protocol describes the evaluation design of the recommissioned Hanen LLLI efficacy trial. A second pilot is currently being carried out by NatCen and Communicate SLT, with a mixed mode of delivery. The main aims of the pilot are to understand the mixed mode of delivery, staff perceptions, challenges and facilitators to delivery and participation, and to assess how learning from Hanen LLLI is cascaded within settings.

Intervention

Learning Language and Loving It™ - The Hanen Program® for Early Childhood Educators (Hanen LLLI) is a training program for EY practitioners to promote social, language and literacy learning in nurseries. It is a CPD program designed to provide staff with practical strategies to enhance children's communication and language skills through specialised ways of interacting and communicating with children during normal daily routines. Hanen LLLI was developed by the [Hanen Centre](#)⁴, in Canada and has not been widely used in the UK.

Updated logic model

This evaluation is based on a logic model. An intervention logic model had been initially developed for Hanen LLLI 1 in conjunction with Communicate SLT, and this has been further updated for this trial during the Theory of Change (ToC) workshop. The updated logic model can be found in **Appendix 1: Intervention logic model** It outlines the inputs provided by Communicate SLT and nursery settings and the sequence of activities implemented by Communicate SLT, Program Leaders, nursery staff and settings. It sets out the intended impacts of the programme for local areas, settings and children and the short, medium and

² For more information on the process of collecting data for the pilot see the pilot study plan [here](#).

³ The protocol for the initial efficacy trial of Hanen LLLI can be found [here](#), while the Statistical Analysis Plan can be found [here](#).)

⁴ The Hanen Centre's mission is to enable parents and professionals to transform their daily interactions with young children to build the best possible lifelong social, language and literacy skills

long-term outcomes for staff, settings and children that are expected to lead to these impacts.⁵ Some of the main changes that have been made to the logic model since Hanen LLLI 1 logic model include:

- **Inputs:** the addition of tech equipment for the online workshops and video feedback sessions in the programme inputs.
- **Activities:** the combination of in-person and online workshops and video feedback sessions; and the availability of online resources for staff.
- **Outputs:** the inclusion of Private, Voluntary and Independent (PVI) settings.
- **Outcomes:** the inclusion of additional outcomes for setting staff:
 - medium-term outcomes around i. increase in early identification of language delays, ii. increase in practitioners' interactions with children with possible delays, and iii. interventions put in place earlier;
 - long-term outcomes about i. increased autonomy in engaging with children and ii. improved job satisfaction.

The trial logic model will not be updated after completion of the Hanen LLLI pilot. However, any updates based on the pilot findings will be reflected in the trial research questions.

Intervention delivery

In this evaluation, the intervention will be coordinated and delivered by Communicate SLT CIC, a Community Interest Company (CIC) who provide Speech and Language Therapy (SLT) services, based in the North West of England. Communicate SLT are Hanen-certified trainers for some of the Hanen programs but are otherwise not affiliated in any way with The Hanen Centre.

The trial is taking place in three Regional School Commissioner areas: The North (covering Cumbria, the North East and North Yorkshire), East Midlands and the Humber, and the West Midlands.

A key change from Hanen LLLI 1 is that the current trial will take place in both school-based maintained settings and settings from the PVI sector. This addition is motivated by the fact that many children from disadvantaged background receive EY education in PVI settings, but the available evidence base for this sector has typically been weaker in comparison to maintained nurseries. To address this gap, the EEF have been seeking to broaden their engagement with PVI settings. Communicate SLT have prior experience of working in the PVI sector and have included PVI settings in the Hanen LLLI pilot to further explore the best approach of engaging and delivering the intervention in these settings.

There are important contextual differences between school-based maintained settings and PVI settings. For instance, pay among staff tends to be lower at PVI settings compared to maintained settings (EPI, 2018, 2020). The quality of provision and levels of staff qualifications in PVI settings tend to be lower than in maintained settings, with many nursery practitioners in PVI settings not having qualifications beyond Level 2 (EEF, 2018b). And relatedly, PVI

⁵ The logic model does not include contextual variables and causal mechanisms. These will be explored as part of IPE activities.

settings find it more difficult to afford, recruit and retain higher qualified staff (The Sutton Trust, 2020).

Mode of delivery

Delivery of Hanen LLLI 1 is scheduled across 31 weeks. This will include one introductory workshop⁶ to explain the intervention and evaluation, eight training workshops lasting 2.5 hours each, and six individual video feedback sessions per participant. In addition to this, it will include a pre-intervention (baseline) video and a post-intervention (endline) video⁷ of setting staff recording their practice with children.

Delivery of all sessions in Hanen LLLI 1 was intended to be completed in person. Following the disruption caused by COVID-19, the current trial has been adapted to include both in-person and online delivery. The pilot, which will be completed in September 2022, will inform which activities will be delivered in person and which online. At the time of writing, the intention is for workshops 1,2,5,6 and 8 to be completed in person, while all video feedback sessions will take place online. The pilot will also offer the opportunity to test whether online forums, such as websites and social media, can be used to effectively connect and engage practitioners across the programme.

The training workshops are delivered to groups of 10-20 practitioners by qualified and Hanen LLLI-certified speech and language therapists (SLTs) and Early Years Teachers. They are known as Program Leaders, are fully qualified in the UK and are certified Learning Language and Loving It trainers (LLLI trainer) accredited to deliver the Hanen LLLI programme. The individual video feedback sessions are also led by Program Leaders and take a maximum of 45 minutes.⁸ They take place one-to-one with practitioners and will take place online. They focus on guided reflection, with Program Leaders providing feedback on videotaped interactions between EY practitioners and children. Alongside the individual video feedback sessions, practitioners will record their practice by completing a baseline and endline video, followed by a discussion with the Program Leader. The baseline video will establish the extent to which practitioners already make use of Hanen LLLI strategies and identify particular areas of focus. The baseline videos will provide an opportunity for practitioners to reflect on their practice over the course of the intervention and identify key learning for future practice.

Cascading of learning forms another key activity of the Hanen LLLI programme. The formal cascading of training content is enabled through the use of strategies and provision of information which enhances buy-in from senior leadership and participants. The expectations regarding cascading activities increase as the participants develop through the programme. Informal cascading is enabled through the mixing of trained and non-trained practitioners and encouraging information sharing between these two groups. In addition, participating nurseries are provided a wide range of materials to disseminate learning, encourage reflective practice and celebrate achievements and progress.

The intervention activities are prescribed by The Hanen Centre and non-specified adaptations of the course and training materials and handouts are neither allowed nor encouraged. Minor accepted deviations are detailed in the course handbook for Program Leaders.

Intervention content

⁶ The Hanen Centre uses the term 'orientation meetings' instead of 'introductory workshops'.

⁷ Completion of the post-intervention video is not a requirement for receiving an end-of-programme certificate.

⁸ The Hanen Centre prescribe 30-40 minutes per individual video feedback session.

Program Leaders are expected to help facilitate the four broad aims of Hanen LLLI:

- **Education:** provide practitioners with information on language, social and literacy development and on how best to promote these during everyday play activities, conversations and daily routines.
- **Application:** provide practitioners with opportunities to practice and apply strategies and approaches which promote children's development, with feedback from the Hanen LLLI Program Leader.
- **Collaboration:** work together with practitioners as they plan and implement individual programmes for children with specific needs.
- **Peer support:** give practitioners the opportunity to share ideas, issues, and concerns with their colleagues.

Trainers are instructed to use the 4P teaching cycle while providing training:

1. **Prepare:** Give practitioners a reason for learning by starting with asking them to think about what a particular topic means to them in order to tap into personal experience and interest.
2. **Present:** Present facts and information to deepen or expand knowledge in ways that are interesting, interactive, relevant and enjoyable.
3. **Practice:** Create opportunities for practitioners to practice newly learned skills in a variety of hands-on ways with guidance and feedback.
4. **Personalize:** Provide opportunities for practitioners to apply and integrate information into their own situation and to generalise into a variety of situations.

During the programme, practitioners learn practical strategies for engaging with children to enhance their language development, including for example:

- 'OWL' - observe, waiting and listening, rather than asking questions.
- Using a variety of words and modelling extended language.
- Providing opportunities for children to initiate conversation.
- Engaging 'reluctant' children in small groups.
- Tailoring language and approach to match children's styles and skills.

Impact evaluation

Research questions

The research questions which the impact evaluation of Hanen LLLI aims to answer are the following:

1. To what extent did Hanen LLLI lead to changes in children's receptive language outcome as measured by the BPVS? (Primary outcome)
2. To what extent did Hanen LLLI lead to changes in children's receptive and expressive language outcomes as measured by the RAPT? (Secondary outcome)

3. To what extent did Hanen LLLI lead to changes in children’s behavioural outcomes as measured by the SDQ-T? (Secondary outcome)
4. To what extent did Hanen LLLI lead to changes in receptive language as measured by the BPVS for children who are entitled to Early Years Pupil Premium? (Subgroup analysis)
5. To what extent did Hanen LLLI lead to changes in receptive language for lower and higher ability pupils based on the BPVS? (Subgroup analysis)

Design

The evaluation of Hanen LLLI will be designed as a two-arm cluster randomised controlled efficacy trial, with settings as the unit of randomisation and pupils as the unit of analysis. In total up to 165 settings will be recruited. We will randomise settings within regions and by setting type (maintained vs PVI settings). Within recruited settings, up to 17 eligible children will be selected to take part in baseline and outcome testing. The exclusion criteria for settings and pupils in the trial are outlined in the Participants section. All settings signing up will have a 50:50 chance of being assigned to the treatment or control group within each regional-setting type strata. Randomisation, if conducted correctly, should result in no important differences between treatment and control groups in the main determinants of our outcomes of interest (see the Randomisation section for more information on the assignment process). As a result, any discrepancy in outcomes at the end of the trial can be attributed to the intervention itself.

Settings allocated to the treatment group are offered the Hanen LLLI training. Settings assigned to the control condition will implement a business-as-usual approach to language teaching. As an incentive for participation in data collection activities, control settings will receive two payments: a payment of £100 for completion of baseline testing in November 2022 and £900 in August 2023 for completion of endline testing and for completing and sending videos to monitor changes in staff’s practice. The incentive is intended to mitigate the risk that settings are approached about the trial but choose not to participate once assigned to the control group.

The selected outcome measures are age-appropriate, fit well with the Hanen LLLI logic model and were selected in collaboration with the delivery team. The primary outcome of interest is receptive English language as measured by the BPVS-3 age standardised score. The first secondary outcome provides a further measure of receptive and expressive English Language, as measured by the RAPT. In addition to language, the evaluation will also assess differences in socio-emotional development using the SDQ-T.

Table 1: Trial design

Trial design, including number of arms	Two-armed cluster randomised trial
Unit of randomisation	Setting
Stratification variables (if applicable)	Geographic region
variable	Receptive language

Primary outcome	measure (instrument, scale, source)	BPVS-3 age standardised score, 85-115, GL Assessment
Secondary outcome(s)	variable(s)	1. Expressive and receptive language 2. Socio-emotional development
	measure(s) (instrument, scale, source)	1. RAPT 5 th Edition Information and Grammar 2020 raw score, 0-41 Information and 0-39 Grammar, Routledge 2. SDQ-T score, 0-40, Goodman
Baseline for primary outcome	variable	Receptive language
	measure (instrument, scale, source)	BPVS-3 age-standardised score, 85-115, GL Assessment
Baseline for secondary outcome	variable	1. Expressive and receptive language 2. Socio-emotional development
	measure (instrument, scale, source)	1. BPVS-3 age standardised score, 85-115, GL Assessment 2. Not applicable

Randomisation

The unit of randomisation in this trial is the setting-level, with stratification according to region and setting type (mainstream / PVI). Settings that agree to participate in the trial will be allocated to one of the two groups with a 50:50 ratio of settings between treatment and control. Six strata will be constructed from the three Regional School Commissioner areas where the intervention is taking place (The North, East Midlands and the Humber, and West Midlands), and based on setting type (we will distinguish between mainstream and PVI settings). Stratification will ensure that settings from the same region as well as settings with similar type of provision will be evenly allocated to the treatment and control group.

Random allocation of settings to the treatment and control condition will take place in the last week of August and will include all settings that have signed Memoranda of Understanding (MoUs). We will randomise before baseline data collection to facilitate preparations for training delivery undertaken by the developer. To achieve balance in terms of the stratifying measures across the entire sample of settings that expressed interest in this trial, we will randomise all settings at the same time i.e. in the last week of August 2022. All settings will be notified of their allocation following baseline data collection in October (i.e. to accommodate different school timelines, settings in Leicestershire will be notified of their allocation on 13 October, while the remaining settings will be notified on 20 October 2022). We recognise that randomising before baseline data collection is completed presents an additional risk of programme withdrawal. To avoid a substantial decrease in study power (i.e. loss of participants between randomisation and baseline data collection and between baseline and endline data collection) Communicate SLT are recruiting more settings than in Hanen LLLI 1, where they recruited 147 settings. This will create a buffer for school withdrawal between randomisation and baseline data collection or school attrition between baseline and endline data collection. The power calculations assume an overall attrition rate of 20%.

Randomisation will be undertaken in Stata and both do and log files will be used to record the randomisation process. At time of randomisation, the researchers will be blinded to settings'

identity. Setting identifiers will then be merged with group allocation data after randomisation. Having unequal numbers of settings from one region or from one type would mean that there is a high probability that the treatment or control group will be of unequal size. To deal with unequal allocation, we will use the command `randtreat` and the option `misfits(global)` in Stata.

Participants

Setting and practitioner eligibility criteria

Settings will be recruited from across the three Regional School Commissioner areas according to the following eligibility criteria:

- a. Both maintained settings and settings from the PVI sector are eligible for inclusion.
- b. Settings will be recommended to sign-up two-thirds of staff working with 3-4-year-olds to take part in Hanen LLLI. At a minimum, at least 50 per cent of eligible practitioners at each setting (including a teacher or senior member of staff) should be able to participate.⁹
- c. No more than 50 percent of practitioners in a setting should have previously participated in a similar Hanen intervention called 'Teacher Talk' and none should have previously participated in Hanen LLLI.
- d. Settings should ideally have at least fifteen registered 3-4-year-olds to be included in the trial sample. However, we will also consider settings with a minimum of twelve 3-4-year-olds for inclusion.

The inclusion of PVI settings in this trial marks an important change compared to Hanen LLLI 1, where only maintained settings were eligible. The inclusion of PVI settings introduces some important considerations for the trial, which are discussed later in this section and in the Sample size calculations section.

The minimum requirement on the number of 3-4-year-olds per setting is designed to ensure that the trial includes enough pupils to attain the desired level of statistical power. This is discussed further in the Sample size calculations section below. As in Hanen LLLI 1, we will be prepared to include some smaller settings (with a minimum of 12 pupils). This will help to make sure that smaller settings (for example, those in more rural areas of Cumbria) can also be represented in the trial.

The requirement that practitioners should not have previously taken part in Hanen LLLI, and that no more than 50% should have taken part in Teacher Talk, will help to make sure that the trial cleanly captures the impact of Hanen delivery in 2022-23. Beyond these requirements, managers in settings will be able to select practitioners to take part on any basis.

The recruitment of eligible settings is described below. The target sample sizes are described in the Sample size calculations section.

Selection of pupils

We will carry out assessments with up to 17 pupils per setting. If settings have fewer than 17 pupils, we will aim to collect assessment data from all pupils. If settings have more than 17 pupils, we will randomly select 17 pupils for baseline assessments. We will also randomly

⁹ This criterion will be reviewed in light of the findings from the pilot and following recruitment in the trial

select an additional three pupils (or up-to-three, depending on the size of the setting), to act as replacements for any sampled pupils who are absent during the baseline testing. We will carry out the endline assessments with the same pupils who completed the assessments at baseline. Please see the Sample size calculations section below for more information about how the pupil-level sample size was determined.

We will ask participating settings to enumerate all 3 to 4-year-olds for whom consent to participate has been received from their parent or caregiver. This will serve three purposes: (1) to facilitate the random selection of seventeen pupils for baseline assessment, (2) to facilitate longer-term linkage of pupil assessment data with the National Pupil Database (NPD), and (3) to gather information on Early Years Pupil Premium (EYPP) status for the purpose of sub-group analysis. The pupil enumeration forms will collect information on pupils' first and last names, EYPP status date of birth and Unique Pupil Number (UPN). As not all pupils in PVI settings will have a UPN, we will also collect home postcode to help ensure that linkage with the NPD is possible¹⁰.

The pupil enumeration forms will be filled in as electronic spreadsheets in September 2022 and shared with NatCen using a secure upload platform.

Participant recruitment

Communicate SLT will lead on the recruitment of eligible nurseries between February and July 2022. NatCen have supported this process by helping confirm the eligibility criteria and developing information sheets to communicate details about what trial participation will involve.

We will invite all participating settings to complete an MoU in the summer term 2022. This indicates institutional consent to be involved in the study. A research information sheet and privacy notice will be supplied by NatCen explaining that settings' access to the programme will be allocated at random and that participation in the evaluation will involve the collection and processing of children's personal data. At the beginning of the autumn term 2023, participating settings will be required to inform parents/carers of all eligible 3-4-year-olds about their setting's participation in the trial by letter. Parents/carers will be given the opportunity to withdraw their child from the evaluation by contacting their child's setting or the NatCen evaluation team.

Settings will also be asked to identify which staff would take part in the programme and to carry out the pupil enumeration described above. This will happen before the actual treatment allocation is revealed.

Sample size calculations

We will aim to recruit 165 settings in the trial, with half randomly allocated to the intervention condition (in which nurseries receive the Hanen LLLI intervention) and the other half randomly allocated to the 'business as usual' control condition. We have agreed with Communicate SLT that most of the recruited settings will be maintained nurseries (between 60% to 70%). We expect the remaining recruited settings to be PVIs (i.e. 30% to 40%). Within each setting we will sample up-to 17 pupils to complete assessments.

Table 2 shows the estimated Minimum Detectable Effect Size (MDES) for the trial after accounting for expected attrition at both the setting and pupil-level. Our power calculations are

¹⁰ Advice received from the DfE Data sharing team (November 2021) indicates that the best linkage variables for pupils without a UPN would be surname, forename, date of birth and postcode.

informed by Hanen LLLI 1. We assume there is a high pupil-level correlation between baseline and follow-up (0.60) and moderate setting-level correlation (0.36).¹¹ There is limited information available about likely Intra-Cluster Correlations (ICC) in EY settings, but we have based our assumption of 0.185 on what we found in the baseline data collected for Hanen LLLI 1. We use a type-one error rate of 0.05 and a type-two error rate of 0.20 (power of 0.80). We have conducted these power calculations using PowerUp! (Dong and Maynard, 2013).

Table 2: Sample size calculations

			OVERALL	EYPP
Minimum Detectable Effect Size (MDES)			0.217	0.322
Pre-test/ post-test correlations		level 1 (pupil)	0.60	0.60
		level 2 (setting)	0.36	0.36
Intracluster correlations (ICCs)		level 2 (setting)	0.185	0.185
Alpha			0.05	0.05
Power			0.8	0.8
One-sided or two-sided?			2	2
Average cluster size			15	2
Number of settings		Intervention	66	66
		Control	66	66
		Total	132	132
Number of pupils		Intervention	990	132
		Control	990	132
		Total	1980	264

Notes: Power calculations were performed using PowerUp! under an alpha level of 0.05 and power of 0.8. The calculations include estimates of the proportion of variance explained through the included covariates at each of these levels (also known as R^2). The R^2 values here have been estimated by squaring the pre-test post-test correlation. An R^2 value of 0.36 at pupil-level and 0.13 at setting-level is used in the power calculations. We note that EEF protocols usually include sub-group analysis by Free School Meal (FSM) status. Here we use EYPP here rather than FSM, as EYPP data is directly available from settings, whereas FSM is not. EYPP provides schools with additional funding for all 3 to 4-year-olds from low income families. All EYPP are also eligible for FSM.

Attrition is a risk for any panel study and is a key consideration for this trial. We have deliberately powered the trial to withstand a reasonable degree of attrition, based on learning from Hanen LLLI 1 and our understanding of the potential risks of including PVI settings. During Hanen LLLI 1, 16 settings withdrew from the trial before the baseline assessments and a further 22 settings withdrew afterwards. The onset of COVID-19 was a key reason for much of the attrition that took place after the baseline, and we do not expect the pandemic to have such a substantial impact on the current trial. Nevertheless, our experiences from Hanen LLLI 1 indicate that we should plan for a moderate amount of setting-level attrition this time. The current trial also faces an additional risk of attrition due to the inclusion of PVI settings.

¹¹ Power calculations use variance explained by covariates (R^2) as opposed to pre-post-test correlations. We approximate the pre-post-correlation by taking the square root of the R^2 .

Previous studies have shown that the rate of attrition is typically higher in the PVI sector compared to maintained nurseries, with estimates ranging from around 15% to as much as 36%. These high estimates of attrition have informed our decision to include a higher proportion of maintained nurseries in the trial sample than PVIs in order to help ensure that the trial adequately covers the PVI sector, while still remaining resilient to attrition overall.

We also anticipate that the average number of pupils we collect baseline and endline assessment data for will be less than 17 per setting. This is partly because of some expected pupil-level attrition between baseline and endline (for example, some pupils may be absent for the endline testing or may have moved setting after the baseline). A second reason is that we are expecting to include some settings with fewer than 17 pupils in total. As described above, this will allow us the flexibility to include smaller settings that may have as little as 12 pupils.

Given our expectations about attrition, Table 2 is based on the assumption that 132 settings will remain in the trial at endline (representing roughly 20% attrition across the same of recruited settings), with an average of 15 pupils per setting (accounting for 12% attrition at the pupil level). Under these assumptions we would achieve an MDES of about 0.217. This MDES is less favourable than the 0.2 that would be required for the trial to achieve a '5-padlock' security rating. It is possible that we will be able to improve on this in practice if we are able to minimise attrition so that it is smaller than our current assumptions. We will employ a range of strategies to try and keep attrition to a minimum. These include the following:

- We will develop recruitment information materials that are as clear and accessible as possible, whilst also meeting all our obligations around providing the required information about what study participation will involve.
- As outlined in the Randomisation section above, we will overrecruit to account for any withdrawals between randomisation and baseline data collection and attrition between baseline and endline data collection.

We have also conducted power calculations to assess the MDES for subgroup analysis according to pupils EYPP eligibility status. We assume that 13.0% of pupils are eligible for EYPP ([DfE, 2021a](#)). This indicates an MDES of around 0.322.

Outcome measures¹²

Primary Outcome

The primary outcome of interest is English language ability, measured using the British Picture Vocabulary Scale third edition (BPVS-3). The same outcome will also be used as a measure of baseline attainment.

BPVS is an individually administered, norm-referenced, test of receptive vocabulary for Standard English. This test is suitable for children with learning and communication difficulties. A child's receptive vocabulary is tested by asking children to identify pictures that illustrate a given word's meaning.

The BPVS-3 will be collected at baseline in October 2022 and at endline towards the end of the summer term 2023. As outlined in the section above, at baseline settings will have the choice of either having an external assessor visit the setting to carry out assessments or

¹² Please see the [Statistical Analysis Guidance](#).

having setting staff carry out the assessments directly. At endline, all assessments will be done by SLTs.

The primary outcome will be the BPVS-3 age standardised score used to compare children's language ability against a large nationally representative sample of children of similar age. The score is standardised to a mean of 100 indicating whether a child is above or below the BPVS's national standardisations sample. Conversion tables used to standardise the raw BPVS scores are available in the BPVS3 Manual.

The BPVS has several advantages. One advantage is that it is quick to administer and score. It also has a high construct validity, meaning that it is considered a reliable measure of language ability. However, the simplicity and convenience of the assessment also presents limitations. The test is only suitable for measuring hearing vocabulary at a particular point in time. For this reason, is advised to avoid over-generalising conclusions on learning outcomes (Dunn et. al. 2009).

We previously administered BPVS baseline data collection for Hanen LLLI 1 in October 2019 (the trial was cancelled before the endline took place). In Hanen LLLI, we recruited SLTs to carry out the majority of pupil assessments. A small number were undertaken by NatCen researchers due to difficulties in SLT recruitment. Inspection of the old baseline data shows that the distribution of pupil scores was low compared to the reference sample against which the BPVS scores were standardised¹³. This may reflect high levels of need for the Hanen LLLI intervention among settings that were recruited. It is also possible that some improvements may be needed in our briefing materials and guidance for assessors to ensure that assessments are carried out in accordance with the BPVS protocol. We will investigate patterns in the baseline data further in the preparation phase for the new round of baseline data collection. We will not be engaging any NatCen interviewers to conduct assessments in the new trial, which may have been associated with some of the lower-than-expected scores in the previous baseline.

Secondary Outcomes

There are two secondary outcomes for this trial i.e. expressive and receptive language and socio-emotional development. These are discussed further below.

RECEPTIVE AND EXPRESSIVE LANGUAGE

The Renfrew Action Picture Test (RAPT) will be collected at endline as a secondary measure of receptive and expressive language. The RAPT comprises 10 pictures depicting various scenarios. Children's receptive and expressive language is tested by asking children to describe the pictures that they are shown. Children's answers are recorded and then scored according to two separate perspectives: information and grammar.

Each question follows specific scoring guidelines. The raw score for information ranges from 0 to 41, while the raw score for grammar ranges from 0 to 39.

The main advantage of the RAPT is that it provides a snapshot of a child's level of expressive language. However, as with the BPVS, the results represent a snapshot that should be interpreted as part of a wider assessment of language abilities.

¹³ 14.1% of the baseline sample had BPVS scores below the minimum expected for their age group. This means that standardised scores were not available for these pupils, because too few pupils in the standardisation sample obtained scores in that range to enable a standardised score to be constructed.

We will use the raw RAPT score because the test has been standardised only on children who speak English as a first language, so the standardised score will be misleading if applied to children who speak English as an additional language (EAL).

The RAPT will be collected by SLTs at endline. SLTs will not be directly informed of settings' treatment allocation.

SOCIO-EMOTIONAL DEVELOPMENT

The trial will also collect the teacher version of the SDQ as an additional secondary outcome at endline. The SDQ is a brief emotional and behavioural screening measure. The SDQ comprises 25 items. The questionnaire is divided into five subscales measuring emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems and prosocial behaviour.

The SDQ is scored on a total difficulties score, which is generated by summing scores from all the subscales except the prosocial subscale and ranges from 0 to 40. Scores are classified according to four categories: 'close to average', 'slightly raised', 'high' and 'very high', which were defined based on a population-based UK survey¹⁴. This trial will use the SDQ raw scores.

The SDQ has good concurrent validity and discriminant validity and moderate retest reliability (Lundh, Wangby-Lundh & Bjarehed, 2008; Muris, Meesters & van den Berg, 2003; Yao et al., 2009). However, research on the reliability of SDQ has produced mixed results, with some studies reporting satisfactory internal consistency and others highlighting low internal consistency of the subscales (Goodman, 2001; Mieloo et al., 2012).

The assessment will be collected by setting staff that know the children well.

Analysis

PRIMARY OUTCOME ANALYSIS

We will use an Intention-to-Treat (ITT) approach to estimate the impact of Hanen LLLI on the primary outcome. The trial is designed as a two-armed cluster randomised efficacy trial with pupils (level one) clustered within settings (level two). To account for the clustering of pupils within settings, the impact will be estimated using a two-level linear regression model. The BPVS raw score at follow-up will be the dependent variable, with a binary indicator of treatment allocation, baseline BPVS score and geographic region (the randomisation strata) included as independent variables. School-level random effects will be included in the model by allowing the intercept to vary randomly across schools.

Our model will follow EEF statistical analysis guidance (EEF, 2018a). Effects will be presented as *Hedge's g* effect sizes, accounting for clustering of pupils within settings (Hedges, 2007).

SECONDARY OUTCOME ANALYSIS

We will also estimate the impact of Hanen LLLI on two secondary outcomes: the RAPT and SDQ-T.

The analytical approach will be analogous to the primary outcome estimation. We will use an ITT approach, consisting of a multi-level linear regression model with pupils nested within settings. In the case of the RAPT analysis, we will include baseline BPVS scores as an

¹⁴ See the SDQ scoring guidelines on the Youth In Mind website here: <https://www.sdqinfo.org/py/sdqinfo/c0.py>

additional covariate (since we will not have a baseline measure for the RAPT, but baseline BPVS scores are expected to have predictive power for the RAPT at follow-up). We will omit this covariate from the SDQ-T outcome model.

SUBGROUP ANALYSIS

The following subgroup analysis will be undertaken:

1. Children who are registered as eligible for EYPP based on information provided by the settings. We will explore differential effects for EYPP children as they are considered a key target group by the EEF.
2. Children with lower reading ability who scored below a threshold on the BPVS assessment at baseline. Low language achievers are children scoring in the bottom quartile on the age standardised BPVS assessment. A binary indicator (below or above the threshold) will be constructed to define low language achievers. The binary indicator will be defined based on the sample-specific BPVS scores. This will inform if there are any differences in Hanen LLLI for lower language ability children.

We will undertake the subgroup analysis for the primary outcome measure only using the same model specified for the primary outcome.

Compliance

The main framework of analysis for this trial is ITT (for more information see the Analysis section). However, we will also explore programme effects separately for settings that will be allocated to the intervention group that implemented the intervention as intended, based on a compliance score. Compliance is defined as the fulfilment of a set of minimum criteria which determine whether a setting has delivered the Hanen LLLI approach as intended. This will be a binary measure, indicating whether a setting was compliant or not.

The exact criteria on which it will be determined whether a setting is compliant or not will be determined in collaboration with the delivery team, and it will be outlined in more detail together with the compliance analysis approach in the Statistical Analysis Plan (SAP).

At the time of this writing we consider staff attendance at Hanen LLLI training as the main element for compliance. The criteria around training attendance that are considered important are outlined below. For a setting to be deemed compliant they need to fulfil the criteria (items 1 through 4 in **Error! Reference source not found.**) below. The compliance indicators suggest the minimum amount of compliance needed to generate a treatment effect, and they were defined in coordination with the delivery team. It is important to note that the compliance measure at the time of the writing does not monitor if staff in the setting was replaced following training (e.g. prior delivery or during delivery). The process for handling staff replacement will be defined at the SAP stage.

Attendance at the information event and training will be captured via templates designed by NatCen and completed by Communicate SLT.

Table 3: Compliance measure development

Number	Compliance criterion	Data source	Compliance indicator
1	Attendance at information event	MoU summary template	At least 1 member of staff attends so that they understand the commitment

2	Attendance at introductory workshop	Attendance register	At least 1 member of staff per setting attends
3	Attendance of training workshops	Attendance register	At least 1 member of staff attends 6 or more workshops. Workshop 1, 2 and 5 are essential (i.e. if an essential workshop is missed then a setting will not be deemed compliant)
4	Attendance of video feedback sessions	Attendance register	At least 1 member of staff attends 4 or more video feedback sessions

Missing data on any criteria will be scored as zero i.e. the setting will be considered as non-compliant.

In a situation of imperfect compliance (i.e. if some settings are non-compliant), we will undertake a Complier Average Causal Effect (CACE) analysis, by drawing on an instrumental variable (IV) approach, and using a two-stage least squares (2SLS) estimation approach to recover the treatment effect for those who complied with assignment. The first stage estimates if the assignment to Hanen LLLI pushes settings to take up treatment (the first stage regresses treatment assignment on compliance (as defined above)). This provides an estimate of the compliance rate. Results for the first stage will report the correlation between the instrument and the endogenous variable and an F test. The second stage of the IV estimation predicts the outcome using the compliance rate estimated in the first regression by substituting the treatment indicator (assignment to Hanen LLLI) with the compliance rate. The results of this model will answer the research question: ‘To what extent does compliance with the Hanen LLLI delivery requirements lead to improved language outcomes for children?’. This model will be estimated for the primary outcome measure only i.e. for receptive language skills.

Longitudinal follow-ups

Long term follow-up for this trial is currently not considered and thus not discussed further in this protocol.

Implementation and process evaluation¹⁵

The IPE is designed as being complementary to the impact evaluation as it will assist in contextualising and improve our understanding of the impact (or lack thereof) of the programme. Analysis of the two strands will be carried out independently of each other in order to avoid one set of findings influencing analysis or interpretation of the other.

By collecting data on the different components of the logic model through the IPE activities (outlined below), we will seek to provide explanations for the impact evaluation research questions (see the Impact evaluation section for an overview of research questions), specifically by probing participants why and how perceived changes were observed. We will also aim to provide explanations for the two research questions covering subgroup analysis; that is, why Hanen LLLI did or did not lead to i. changes in receptive language as measured

¹⁵ Please follow the principles detailed in the [Implementation and Process Evaluation Guidance \(2019\)](#).

by the BPVS for children who are entitled to Early Years Pupil Premium; and ii. changes in receptive language for lower and higher ability pupils based on the BPVS.

We have adapted our IPE design to incorporate lessons learned from Hanen LLLI 1. This includes:

- **Covering two additional outcomes** in the interviews and survey with practitioners (change to practice and children's social and emotional development).
- **Capturing the perceived outcomes** of the programme on children with language delay via the interviews and survey with practitioners.
- **Administering a survey of staff who did not attend training** to find out more about cascading of learning.
- **Reviewing videos at baseline and endline** to assess changes in the practices of staff who have taken part in Hanen LLLI and those who have not taken part.
- **Conducting paired interviews** with Program Leaders.

Research questions

A process study will be carried out alongside the impact study and aims to answer the following questions:

1. How is Hanen LLLI delivered, and what are the facilitators and barriers to delivery in maintained nurseries and PVI's?
2. What are the perceived benefits of Hanen LLLI for EY staff, nurseries and children?
3. Is there evidence that Hanen LLLI leads to changes in staff practice? How can we better support staff, and assist towards an effective cascading of knowledge?
4. What can be learnt for future delivery of Hanen LLLI?

Research methods

We will conduct a multi-stage process evaluation and take a mixed-method approach to data collection (this will include depth interviews, observations, surveys and the collection of attendance data from the MoU summary template and attendance register). This is designed to collect data from multiple stakeholders (senior staff at treatment and control settings, Hanen-trained and non-trained staff at treatment settings, and Hanen LLLI Program Leaders) at different time-points (pre-intervention, early and ongoing implementation, and post-intervention) and balance depth and breadth of insight. Table 4 presents an overview of the research methods alongside the research questions addressed and theory of change relevance. Further details of the research methods are provided below.

Observations of Hanen LLLI training

We will observe 8 Hanen LLLI workshops across the course of the program. This will aim to increase our understanding of how the training is delivered, including adherence to the Hanen LLLI guidance, resources used, duration of the sessions and any adaptations, as well as staff's engagement with the training and perceived quality. We will observe two in-person workshops and two online workshops at different time-points of programme delivery, with two different

groups of Program Leaders and staff being observed for each workshop (8 in total). The groups to be observed will be selected to cover a range of geographical areas and types of settings (maintained and PVI). Observations will take place between **November 2022 and May 2023**.

Early implementation interviews

We will conduct 10-12 in-depth, semi-structured interviews with senior staff in treatment nurseries at the start of the programme. These will gather contextual information on the settings and usual practice to support children's language and communication development. They will also explore how senior staff plan to support staff to attend Hanen LLLI workshops and any intervention challenges they anticipate. The interviews will also look at whether settings' approaches to language and communication development have changed as a result of the pandemic and how. Settings will be selected based on the type and size. We will also seek to cover a range of geographical areas and where possible a range of setting profiles in terms of children eligible for EYPP and with EAL. The interviews are expected to last around 30 minutes and will take place online / via telephone between **November and December 2022**.

Ongoing delivery interviews

We will conduct 10-12 in-depth, semi-structured interviews with senior staff in treatment nurseries during ongoing delivery. These will explore how the delivery of Hanen LLLI has been progressing, including challenges and facilitators, any cascading of the Hanen LLLI strategies to non-trained staff and the perceived outcomes of the programme for staff, the setting environment and children. Nurseries will be sampled based on the same criteria as the early implementation interviews. Half of these interviews will be carried out with settings that participated in the early implementation interviews to facilitate holistic and longitudinal understanding of programme delivery. The interviews will last approximately 45 minutes and will be conducted online / via telephone between **March and April 2023**.

'Business as usual' interviews

Eight in-depth, semi-structured interviews will be conducted with senior staff at control nurseries. These will gather data on usual practice in nurseries that do not take part in Hanen LLLI, including engagement with any communication and language development interventions. Nurseries will be selected based on type and size of setting, with an aim to achieve a mix of geographical areas. Interviews will last approximately 30 minutes and will be conducted online / via telephone between **January and March 2023**.

IPE site visits

We will select eight settings for site visits, aiming to achieve a range in setting type, size and area. The sample will include at least one (i) large group-based provider to explore Hanen LLLI in the context of large numbers of staff and children; and (ii) setting with a less formal literacy curriculum and/or inexperienced EY staff. The Hanen LLLI 1 IPE suggests that Hanen LLLI is of particular benefit in those settings. The site visits will take place between **January and April 2023**. Fieldwork is expected to be mainly face-to-face, supplemented by online / telephone interviews where necessary. Each site visit will include:

- **Interviews with 1-2 members of staff taking part in Hanen LLLI.** These will explore staff's views on training, barriers and facilitators to delivery, perceived outcomes on practice, adequacy of support and any cascading of learning.

- **Interviews with setting managers** to explore motivations, barriers and facilitators to engagement with Hanen LLLI. The interviews will also look at any adaptations and support required, cascading of knowledge, perceived outcomes on staff and children and how Hanen LLLI interacts with existing practices.
- **Observations of 1-2 video feedback sessions** to understand the content of these sessions, practitioners' responsiveness, and any barriers and facilitators to delivery. The observations will take place online and will only go ahead if practitioners are comfortable with a NatCen researcher observing their online video feedback session.

Paired interviews with Program Leaders

We will conduct in-depth, semi-structured paired interviews with six Program Leaders to provide evidence on training and video feedback delivery. This will include any adaptations to the prescribed Hanen LLLI guidance, perceived acceptability of the programme, perceived outcomes and recommendations for improvements. Interviews will last approximately 45 minutes and will be conducted online / via telephone between **April and May 2023**.

Survey of staff taking part in Hanen LLLI

We will conduct a short post-intervention web survey of all setting staff who took part in Hanen LLLI to gather feedback on the intervention, including engagement, barriers and facilitators and perceived outcomes for staff and children. This will take approximately 15 minutes to complete and will be administered between **June and July 2023**.

Survey of non-Hanen trained staff in treatment nurseries

We will conduct a short post-intervention web survey of all nursery staff in treatment settings who did not take part in Hanen LLLI to get a better understanding of the cascading of learning from the programme. This will take approximately 10 minutes to complete and will be administered in **June-July 2023**.

Videos of staff interacting with children

Two videos will be produced by the same member of staff in both treatment and control settings: one at the beginning and one end of the year, with the aim to capture changes in staff practice. The videos will be reviewed by qualified SLTs and Hanen LLLI licensees who will score staff interactions with children according to the extent to which they implement Hanen LLLI strategies. Pre-intervention and post-intervention scores will be compared to assess changes in the practices of staff who have taken part in Hanen LLLI and those who have not taken part and have continued with usual practice. The video reviewers will not be involved in programme delivery. They will undergo specific training in the completion of the Hanen video analysis tool from a Hanen LLLI licensed trainer. When reviewing videos, they will be blind as to whether the practitioner in the video is in the treatment or control group.

A short web form to gather cost information

We will collect and analyse cost data in line with EEF guidelines¹⁶. At the end of the academic year, NatCen will circulate a web form to project leads at treatment settings to gather information on the costs of taking part in Hanen LLLI, which will be used to calculate a per-pupil cost of the programme. The costing form will be incorporated into the survey of staff

¹⁶ https://d2tic4wvo1iusb.cloudfront.net/documents/evaluation/evaluation-design/Cost_Evaluation_Guidance_2019.12.11.pdf

taking part in Hanen LLLI. Communicate SLT will also be asked to complete a proforma detailing delivery costs. We will estimate the per-pupil cost over three years based on annual average net costs divided by the number of participating pupils.

Table 4: IPE methods overview

Research methods	Data collection methods	Participants/ data sources (type, number)	Data analysis methods	Research questions addressed	Implementation/ logic model relevance
Observations	Semi-structured observations of training workshops throughout the programme (in-person and online)	Staff & Program Leaders/ Training workshops (8)	Thematic analysis	RQ1	Fidelity, quality, adaptation, responsiveness
Observations	Semi-structured observations of video feedback sessions throughout the programme (in person and/or online)	Staff & Program Leaders/ Video feedback sessions (8-16)	Thematic analysis	RQ1	Fidelity, quality, adaptation, responsiveness
Interviews	Semi-structured interviews during early implementation (45 mins, online/via telephone)	Senior staff in treatment settings (10-12)	Thematic analysis	RQ1	Usual practice, fidelity, adaptation, responsiveness, setting conditions
Interviews	Semi-structured interviews during ongoing delivery (45 mins, online/via telephone)	Senior staff in treatment settings (10-12)	Thematic analysis	RQ1, RQ2, RQ3, RQ4	Fidelity, quality, adaptation, responsiveness, perceived outcomes, setting conditions
Interviews	Semi-structured interviews covering business as usual (30 mins, online/via telephone)	Senior staff in control settings (8)	Thematic analysis		Usual practice
Interviews	Semi-structured interviews at mid-point as part of site visits (45 mins, in person or online/via telephone)	Senior setting staff in treatment settings (8)	Thematic analysis	RQ1, RQ2, RQ3, RQ4	Fidelity, quality, adaptation, responsiveness, perceived outcomes
Interviews	Semi-structured interviews at mid-point as part of site visits (45 mins, in person or online/via telephone)	Hanen-trained setting staff (8-16)	Thematic analysis	RQ1, RQ2, RQ3, RQ4	Fidelity, quality, adaptation, responsiveness, perceived outcomes
Interviews	Semi-structured paired interviews towards the end	Program Leaders (6)	Thematic analysis	RQ1, RQ2, RQ4	Fidelity, adaptation, responsiveness,

	of the programme (45 mins, online/via telephone)				perceived outcomes
Survey	Online post-intervention survey (15 mins.)	Hanen-trained setting staff (all)	Descriptive statistics; thematic analysis	RQ1, RQ2, RQ3, RQ4	Usual practice, fidelity, responsiveness, perceived outcomes
Survey	Online post-intervention survey (10 mins.)	Non-Hanen trained staff in treatment settings (all)	Descriptive statistics; thematic analysis	RQ1, RQ2, RQ3	Usual practice, fidelity, responsiveness, perceived outcomes
Observations	Structured observations and scoring of pre- and post-intervention videos of staff interacting with children	Staff in treatment and control settings (150)	Descriptive statistics	RQ3	Changes in staff practice
Survey	Post-intervention web form to collect costs information	Project leads at treatment settings (75), Communicate SLT	Descriptive statistics	RQ5	Cost per-pupil
Attendance register	Register of staff attendance throughout the programme	Setting staff, Communicate SLT	Frequency counts; compliance analysis	RQ1	Compliance, dosage, reach

Analysis

Given that a key focus of the evaluation is to gather evidence against the intervention logic model, the analysis approach is primarily deductive but with enough flexibility to capture emergent and unforeseen themes from the data.

Interviews will be digitally audio-recorded with permission from participants, and professionally transcribed. We will use Framework in NVivo 11, a systematic approach to qualitative data management, developed by NatCen, to chart (collate and summarise) transcribed data by theme and case (Ritchie *et al.* 2013). Using the themes covered in topic guides and new emerging themes, we will assemble a matrix in which each row represents an individual interview and each column a theme and any related sub-themes. We will then summarise the interview data in the matrix, including illustrative verbatim quotes where appropriate.

Analysis of workshop and video feedback sessions will adopt a similar approach. We will record workshop and video feedback sessions using pre-written observation proformas. We will then create different analytical frameworks and a series of matrices in Microsoft Excel, each relating to a different thematic issue. The columns in each matrix will represent the key sub-themes or topics and the rows will represent individual workshop or video feedback sessions. We will then 'transfer' the observation and video feedback session notes to the matrix.

At an early stage of analysis, the IPE research team will chart a selection of transcripts, drawing on the relevant analytical frameworks. The team will then gather to address any inconsistencies and ambiguities and to refine the analytical frameworks and relevant themes and sub-themes.

Once all interviews, workshop and video feedback sessions are coded in the matrix, we will analyse the data. This will involve a phase of 'detection', which includes studying the elements participants said about a given topic, listing these and then sorting them thematically. Once we identify different themes in the data, we will create higher level categories that work as meaningful conceptual groupings for participants' views and experiences.

Quantitative data will be analysed in SPSS using descriptive statistics (frequencies and cross-tabulations) to identify patterns and trends.

Triangulation of all data and thematic synthesis by the main implementation domains will provide a comprehensive assessment of implementation and help to explain the impact evaluation findings. Any convergent and divergent findings with the impact evaluation will be explained with reference to the collected evidence, nature of the programme and intervention context.

Cost evaluation

The costs associated with programme delivery will be estimated in line with the convention outlined in the EEF cost guidelines, which relies on the 'ingredients method' principle (Levin et al., 2018) i.e., we will account for all resources necessary to implement Hanen LLLI, regardless of who incurs the costs. The cost evaluation will estimate average marginal costs per pupil per year over a three years period. The programme costs will be divided into pre-requisites, start-up costs and recurring costs. The estimated costs will be adjusted by the year when they are incurred to account for changes in value-of-money over time and inflation (EEF, 2019).

The costs in this trial will be estimated using two different data sources:

1. Data provided by Communicate SLT on costs related to delivering Hanen LLLI workshops, video feedback sessions and any other costs associated with providing support to settings that are part of usual Hanen practice.
2. Data on the costs of participation that fall on settings, such as the hours of supply and cover needed for staff to attend workshops. This data will be provided by settings directly and will be collected by NatCen using a short web form.

Ethics and registration

Ethical approval for this study was obtained from NatCen's Research Ethics Committee in February 2022. The NatCen REC reviewed the study design to confirm compliance with internal ethical standards.

The trial will be registered with the International Standard Randomised Controlled Trial Number (ISRCTN) in the Spring 2022.

Data protection¹⁷

NatCen will obtain personal data from settings and pupils at baseline and pupil outcome data at endline as a data controller and data processor. Setting and pupil level information will be collected and processed on the 'legitimate interest' basis. NatCen will process the data for the legitimate purpose of conducting the evaluation of Hanen LLLI. No special category data

¹⁷ Please see the [Data Protection Statement](#) for EEF Evaluations.

will be collected as part of the evaluation. Setting staff and pupil parents will be provided with accessible information leaflets and privacy notices that explain the use, storage, and secure handling of the data. Consent will be obtained on opt out basis i.e. participants have an option to withdraw themselves or their child from the study. Parents and staff will receive a hard-copy or e-version of the relevant privacy notice, and the same will also be published on the study website. Participants taking part in IPE activities (interviews, observations, surveys) will be asked to 'opt in' and reminded that participation is voluntary.

NatCen will store and handle all data securely and confidentially in line with the GDPR. Only the research team and approved third parties listed in the privacy statement (e.g. transcription agency, speech and language therapists conducting testing) will have access to the data collected as part of the evaluation. Reports and other publications arising from this research will not identify any individual setting, staff member, or pupil. Settings or individual staff who no longer wish to take part in the evaluation can request to have their data deleted at any point prior to the submission of the draft report and before data archiving takes place.

Data sharing between NatCen and Communicate SLT will be governed by a data sharing agreement (DSA) covering the pilot and trial. The data shared between Communicate SLT and NatCen includes contact details and information on nursery settings recruited for the pilot and trial, staff taking part in Hanen LLLI training and staff attendance; as well as pre-intervention and post-intervention scores from one staff member per participating setting. Only data which is necessary for the evaluation will be shared.

For the purpose of research and archiving, NatCen will share data from the impact evaluation with The Department for Education, the EEF's archive manager, the Office for National Statistics and potentially other research teams. At the end of the research project, this data will be submitted to the ONS, SRS in the EEF data archive (this is managed by FFT). This will include data only identifiable to the Department for Education and no information will be archived that could be used to directly identify individual pupils. Further matching to NPD and other administrative data may take place during later research. All data will be securely deleted from NatCen's network one year after the end of the project.

Personnel

Delivery team

The intervention is managed and implemented by Communicate SLT. They are responsible for recruitment, delivery of workshops, video-feedback sessions and intervention delivery. The delivery team also provided input into the design of the trial i.e. in the selection of outcome measures, and in the development of compliance measure. The delivery team is led by Caroline Coyne, and also involves Helen Laycock. They are supported by Janice Greenberg, a Program Director at the Hanen Centre.

Evaluation team

The evaluation is undertaken independently by education and evaluation specialists at NatCen. The evaluation team is responsible for the outcome and process evaluation, trial design, analysis, and reporting. The evaluation team in NatCen involves: Dr Jonah Bury (overall study lead); Dr Sashka Dimova (impact evaluation lead); Molly Scott (impact evaluation lead researcher); Sehaj Bhatti (project management, and process evaluation); Alessandra Sciarra (impact evaluation support). The evaluation team will also benefit from advice and quality assurance from Daniel Phillips and Dr. Tina Haux.

Risks

Table 5: Risk register

Risk	Likelihood / Impact	Mitigation/Contingency
EY settings or practitioners are reluctant to sign up to the evaluation	Likelihood: Medium Impact: Medium	Recruitment was successful in Hanen LLLI 1, so this risk is relatively low. We will work closely with the developers using our existing Hanen materials to ensure they are up-to-date and available early to aid recruitment.
Difficulty in recruiting sufficient SLTs for pupil assessments	Likelihood: Low Impact: Medium	Learning from Hanen LLLI 1 has been incorporated into our current evaluation design, with dedicated resource to recruit and onboard SLTs from our existing networks. We will start this process early. We also note that the trial is taking place in additional areas, compared to Hanen LLLI 1. This should help make it easier to recruit SLTs across the study regions, as the settings are less geographically concentrated this time around.
EY settings drop out of the evaluation	Likelihood: Medium Impact: Medium	The evaluation has been designed to minimise burden on settings, whilst ensuring sufficient depth of data. We will give settings advance notice of research activities and arrange visits to suit their availability. We will work closely with the developers to address concerns about research burden.
Children's parents or EY settings oppose testing of children	Likelihood: Low Impact: High	Opposition to testing is likely to be a low risk. In Hanen LLLI 1, nearly 90% of children were successfully assessed, and assessors reported that children enjoyed assessments. We will ensure that our communication with settings and parents is clear and informative, to allay potential concerns. We will also incorporate learning from Hanen LLLI 1 as to how to conduct the assessment in the most child-friendly manner possible.
COVID-19 disrupts programme delivery or data collection	Likelihood: Low Impact: High	Hanen LLLI has already adjusted its design to include remote training methods, which can continue in the event of further COVID-19 related restrictions.
Pupil and staff-level attrition	Likelihood: Medium Impact: High	The inclusion of PVLs risks attrition at the pupil and staff-level (based on previous EEF experience). We have accounted for attrition in our power calculations to anticipate this risk.
Evaluation team member absence or turn-over	Likelihood: Medium Impact: Medium	NatCen staff have a three-month notice period to allow sufficient time for handover. The team can be supplemented by researchers with relevant topic and methodological expertise from the larger pool of NatCen researchers.

Timeline

Table 6: Timeline

Dates	Activity	Staff responsible/leading
December 2021-January 2022	Set-up meetings	NatCen
January 2022	IDEA Workshop	NatCen
January-February 2022	Recruiting settings (pilot)	Communicate SLT
February 2022	Ethical review	NatCen
February-March 2022	Logic model review	NatCen
February-July 2022	Pilot programme delivery	Communicate SLT
March 2022	Completion of data sharing agreement	NatCen, Communicate SLT
March-2022	Developing recruitment materials (MoU, setting information sheet, parent information letter)	NatCen
March-July 2022	Recruiting settings	Communicate SLT
April 2022	Completion of trial protocol	NatCen
May 2022	Contract agreement finalised	NatCen
May 2022	Completion of trial registration	NatCen
June 2022	Presentation of pilot interim findings	NatCen
August 2022	Randomisation	NatCen
August 2022	Submission of draft pilot report	NatCen
September 2022	Presentation of pilot findings	NatCen
September 2022	Pupil information to be collected	NatCen
September 2022	Selection of pupils to undertake testing	NatCen
October 2022	Baseline outcome testing	NatCen
November 2022	Submission of SAP	NatCen
November 2022-June 2023	Programme implementation	Communicate SLT
November 2022	Pre-intervention staff videos	Communicate SLT
November 2022-June 2023	Compilation of training attendance records	Communicate SLT
November 2022-May 2023	Workshop observations	NatCen
November-December 2022	Early implementation interviews	NatCen
March-April 2023	Ongoing delivery interviews	NatCen

Dates	Activity	Staff responsible/ leading
January-March 2023	Business as usual interviews	NatCen
January-April 2023	Site visits	NatCen
April-May 2023	Program Leader interviews	NatCen
May-June 2023	Post-intervention videos	Communicate SLT
June-July 2023	Staff surveys and costing form incorporated within the survey of staff taking part in Hanen LLLI	NatCen
June-July 2023	Outcome testing	NatCen
July-October 2023	IE and IPE analysis and reporting	NatCen
October 2023	Submission of draft EEF report	NatCen
January 2024	Final EEF report	NatCen

References

- Cabell, S. Q., Justice, L. M., Piasta, S. B., Curenton, S. M. et al. (2011) 'The Impact of Teacher Responsivity Education on Preschoolers' Language and Literacy Skills' *American Journal of Speech Language Pathology* 20 (4): 315–30.
- Department for Education [DfE] (2018) [Education provision: children under 5 years of age, January 2018](#) London: DfE
- Department for Education [DfE] (2019a) [Early Years Foundation Stage profile results: 2018 to 2019](#) London: DfE
- Department for Education [DfE] (2019b) [Childcare and Early Years providers survey: 2019](#) London: DfE
- Department for Education [DfE] (2021a) [Education provision: children under 5 years of age](#) London: DfE
- Department for Education [DfE] (2021b) [Statutory framework for the early years foundation stage](#) London: DfE
- Department for Education [DfE] (2021c) [Changes to the early years foundation stage \(EYFS\) framework](#) London: DfE
- Dunn, L., Dunn, D., Sewell, J., Styles, B., Brzyska, B., Shamsan, Y., Burge, B. (2009) *The British Picture Vocabulary Scale Third Edition, Manual*. GL Assessment.
- Dong, N. and Maynard, R. (2013) PowerUp!: A Tool for Calculating Minimum Detectable Effect Sizes and Minimum Required Sample Sizes for Experimental and Quasi- Experimental Design Studies, *Journal of Research on Educational Effectiveness*, 6(1), pp24-67, DOI: 10.1080/19345747.2012.673143
- Education Endowment Foundation [EEF] (2018a) *Statistical Analysis Guidance for EEF Evaluations* London: EEF
- Education Endowment Foundation [EEF] (2018b) [Independent Evaluation of the Maths Champions Programme](#) London: EEF
- Education Endowment Foundation [EEF] (2019) *Cost evaluation guidance for EEF Evaluations* London: EEF
- Education Policy Institute [EPI] (2018) [The early years workforce: a fragmented picture](#) London: EPI
- Education Policy Institute [EPI] (2020) [The stability of the early years workforce in England](#) London: EPI

Girolametto, L., Weitzman, E. and Greenberg, J. (2003) 'Training day care staff to facilitate children's language' *American Journal of Speech-Language Pathology* 12: 299– 311.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40 (11), 1337-1345.

Hedges, L. V. (2007) 'Effect Sizes in Cluster-Randomized Designs' *Journal of Educational and Behavioral Statistics* 32(4): 341–370.

Law, J., Charlton, J., Dockrell, J., Gascoigne, M., McKeane, C. and Theakston, A. (2017) [Needs, provision, and intervention for preschool children from socio-economically disadvantaged backgrounds](#) London: Education Endowment Foundation

Levin, H. M. et al., 2018. *Economic Evaluation in Education. Cost-Effectiveness and Benefit-Cost Analysis (Third Edition)*. London: Sage Publishing.

Lundh, L.G., Wangby-Lundh, M., & Bjarehed, J. (2008). Self reported emotional and behavioral problems in Swedish 14 to 15-year-old adolescents: A study with the self-report version of the Strengths and Difficulties Questionnaire. *Scandinavian Journal of Psychology*, 49, 523–532.

Mieloo, C., Raat, H., van Oort, F., Bevaart, F., Vogel, I., Donker, M., & Jansen, W. (2012). Validity and reliability of the strengths and difficulties questionnaire in 5-6 year olds: Differences by gender or parental education. *PLoS One*, 7 (5), 1-8.

Muris, P., Meesters, C., & van den Berg, F. (2003). The Strengths and Difficulties Questionnaire (SDQ): Further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *European Child and Adolescent Psychiatry*, 12 (1), 1–8.

Ofsted (2017) [Bold Beginnings: The Reception curriculum in a sample of good and outstanding primary schools](#) London: Ofsted

Ofsted (2020) [Main findings: childcare providers and inspections as at 31 March 2020](#) London: Ofsted

Piasta, S. B., Justice, L. M., Cabell, S. Q., Wiggins, A. K., Turnbull, K. P. and Curenton, S. M. (2012) 'Impact of professional development on preschool early years practitioners' conversational responsivity and children's linguistic productivity and complexity' *Early Childhood Research Quarterly* 27 (3):. 387–400.

Ritchie J. et al. (2014). *Qualitative research practice: A guide for social scient students and researchers*. London: Sage

The Sutton Trust (2020). [Early years workforce review](#). London: The Sutton Trust.

Yao, S., Zhang, C., Zhu, X., Jing, X., McWhinnie, C. M., & Abela, J. R. Z. (2009). Measuring Adolescent Psychopathology: Psychometric Properties of the Self-Report Strengths and Difficulties Questionnaire in a sample of Chinese adolescents. *Journal of Adolescent Health, 45*, 55–62.

Appendix 1: Intervention logic model

Learning Language and Loving It™ - The Hanen Program® for Early Childhood Educators Logic Model

